

Getting ahead of the market: How big data is transforming real estate

Many real estate firms have long made decisions based on a combination of intuition and traditional, retrospective data. Today, a host of new variables make it possible to paint more vivid pictures of a location's future risks and opportunities.

Gabriel Morgan Asaftei, Sudeep Doshi, John Means, and Aditya Sanghvi



In Boston, the price of homes within a quarter of a mile of a Starbucks jumped by more than 171 percent between 1997 and 2014, 45 percentage points more than all homes in the city, according to a February 2015 report by the brokerage and information website Zillow. Over the past decade, Seattle apartment buildings within a mile of specialty grocery stores like Whole Foods and Trader Joe's appreciated in value faster than others.

While the impact of proximity might be intuitive, home prices are not just driven by having nearby grocery stores. Rather, they are driven by access to the right quantity, mix, and quality of community features. More is not always better; for example, though having two specialty food stores within a quarter of a mile correlates with an increase in property prices, having more than four of them within that same distance correlates with lower prices.

These nonlinear relationships are observed across many American cities. And the sweet-spot intersection of density and proximity to community amenities varies among cities and even neighborhoods, obscured by a growing mass of data that is increasingly difficult to tame.

Power of nontraditional data

In our conversations with developers and investors, we often hear frustration with the disconnect between the availability of data and the difficulty of harnessing it for quick, actionable insights. Developers and investors have always sought to understand where to acquire property and when to trigger development. Portfolio holders need to optimize their holdings and regularly assess conditions that lead them to divest or capture value.

Being slow to identify subtle trends means leaving money on the table. Conversely, being a first mover on a compelling (and perhaps inconspicuous) opportunity translates into significant advantage. Why is it so hard to claim that spot as a first mover?

How can real estate developers and investors keep track of so much data and quickly find hidden patterns—and harness them for profitable investments? And what has prevented them from doing so?

Conventional analytical methods and data sources make it challenging to draw clear hypotheses and build robust business cases. Analysts must sift through tens of millions of records or data points to discern clear patterns and place their bets with few supporting tools to help glean insights from that material. By the time an investor can collect, compile, and process the data needed to distill action, the best opportunities are gone.

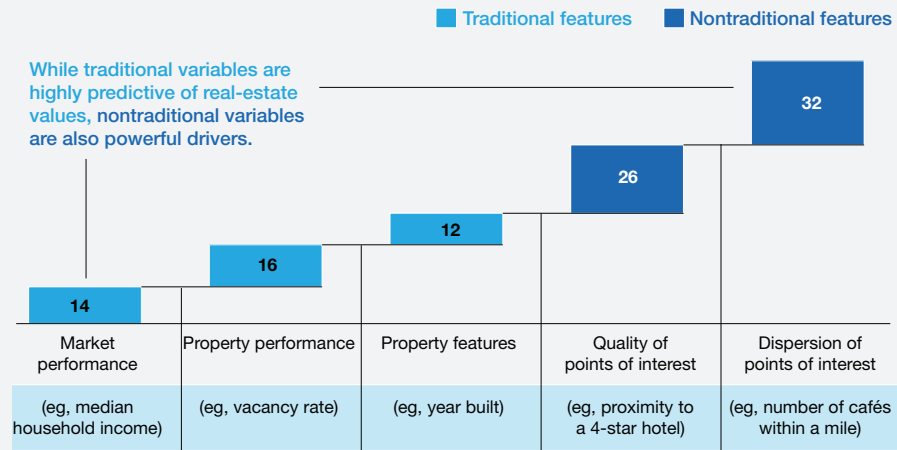
At the same time, new and unconventional data sources are becoming increasingly relevant. Resident surveys, mobile phone signal patterns, and Yelp reviews of local restaurants can help identify “hyperlocal” patterns—granular trends at the city block level rather than at the city level. Macroeconomic and demographic indicators, such as an area's crime rate or median age, also inform long-term market forecasts.

Thousands of nontraditional variables can be linked to diverging, location-specific outcomes (Exhibit 1). These variables include:

- number of permits issued to build swimming pools
- change in number of coffee shops within a one-mile (1.6 km) radius
- building energy consumption relative to other structures in the same zip code
- in-office mobility, based on frequency of elevator movement
- tone of Yelp reviews for nearby businesses

Exhibit 1 Nearly 60 percent of predictive power can come from nontraditional variables.

Proportion of predictive power, % share



McKinsey&Company

This information is not traditionally considered real estate data, but stitching such data points together can more accurately predict hyperlocal areas with outsized potential for price appreciation.

Art of the possible

One way to stitch together the data through advanced analytics is to use machine learning algorithms, which make it significantly easier to aggregate and interpret these disparate sources of data. Technology solutions automate the data collection by accessing application programming interfaces (APIs) and connecting various databases before preparing the data for analysis. After all, it is not the raw data that creates value, but the ability to extract patterns and forecasts and use those predictions to design new market-entry strategies.

Let's say you are a developer who wants to identify underused but high-value parcels zoned for development. Data sources on previous transactions,

such as the Multiple Listing Service, exist and are widely established as the traditional cornerstone of information on both residential and commercial real estate assets. However, these databases have limited value for anticipating future potential, not having been designed for that purpose. Advanced analytics can quickly identify areas of focus, then assess the potential of a given parcel with a predictive lens. A developer can thus quickly access hyperlocal community data, paired with land use data and market forecasts, and select the most relevant neighborhoods and type of buildings for development. Further, that developer can optimize development timing, mix of property uses, and price segmentation to maximize value.

Alternatively, for an asset manager who wants to expand and optimize a portfolio of multifamily buildings, machine learning algorithms can rapidly combine macro and hyperlocal forecasts to prioritize cities and neighborhoods with the highest

demand for multifamily housing. This allows the asset manager to identify buildings in areas that are undervalued but rising in popularity.

Advanced analytics cannot serve as a crystal ball. In most cases, it should only support investment hypotheses, not generate them. But when it comes to these classic real estate conundrums, advanced analytics can rapidly yield powerful input that informs new hypotheses, challenges conventional intuition, and sifts through the noise to identify what matters most.

Impact of a data-driven approach

A successful data-driven approach can yield powerful insights. In one example, an application combining a large database of traditional and nontraditional data was used to forecast the three-year rent per square foot for multifamily buildings in Seattle. These machine-learning models predicted rents with an accuracy rate that exceeded 90 percent.

In addition, the exercise illustrated the power of using nontraditional data. Perhaps unsurprisingly, variables related to traditional data sources—for instance, vacancy rates—correlated with future values. But variables related to nontraditional data, such as proximity to highly rated restaurants or changes in the number of nearby apparel stores, explained 60 percent of the changes in rent.

In accounting for these nontraditional variables, buildings located in the same zip code can have widely disparate outcomes in terms of rental performance (Exhibit 2). Two buildings that are seemingly identical when evaluated by traditional metrics can ultimately experience very different growth trajectories. It is easy to imagine how this disparity at the individual building level, when applied across a series of investments, can drive dramatic results at the portfolio level.

One major advantage of creating applications powered by advanced analytics is their ability to be scaled for use in other scenarios. For example, the same application used to forecast rents can also be used in scenarios such as:

- pressure-testing expectations for individual properties within high-growth markets, aiding in the choice of properties to invest in or places to exit or divest
- identifying individual assets that will hold their value in otherwise declining markets
- making capital expenditure decisions on specific properties (for example, calculating the return on investments or stabilized yield on cost in unit upgrades)
- comparing predictive-model outputs to the forecasts of traditional sources of market information, such as brokers



Making it happen

Building advanced analytics into a portfolio is no straightforward task. Collecting enough data to build accurate algorithms takes time. Manually scrubbing data for use in analytics can be costly. And despite the rise of organizations vying to capture value from advanced analytics across industries, relatively few achieve it at scale.

To face this challenge head-on, companies might begin by employing analytics in executing their most critical strategic imperatives; pursuing data-cleansing efforts based first on the most valuable use cases; and establishing clear processes for data governance, interpretation, and decision making. Ultimately, data analytics should have its own strategic direction with long-term roles and goals beyond just a few pilot projects and use cases.

Exhibit 2 Using **nontraditional metrics**, advanced analytics enables more-accurate predictions about two buildings in the same zip code.

Similar-looking assets in traditional metrics ...

Building 1			Building 2
	\$1.74	Rent per square foot	\$1.70
	1.02:1	Rent per square foot vs zip-code average	1:1
	2%	Vacancy rate	2%
			

... can look substantially different in nontraditional metrics ...

3.28	Average rating of bars within half mile	3.93
3.80 miles	Distance to high-priced restaurant	0.74 miles
16	Number of gas stations within 2 miles	4
20	Number of drinking establishments within 2 miles	5
60	Number of food stores within 2 miles	22

... which ultimately matter more.

\$1.95	3-year predicted rent per square foot, using advanced analytics	\$2.10
\$1.91	Actual rent per square foot	\$2.12
3.9%	Predicted annual growth	7.3%

McKinsey&Company

Developers and investors who want to harness the power of advanced analytics face an additional constraint: few job candidates have the ability to understand business goals, write code and algorithms for analysis, develop decision-making tools, and clearly translate these elements into

advice for business leaders. That is why the most effective organizations build teams that include a variety of experts—scientists who build models, engineers who create data architecture, and translators who help bridge technical expertise with strategic business goals or actions.

When it comes to attracting people with these skills, salary is undoubtedly a factor. But discussions with data professionals uncover something fundamentally more important: the most talented employees do not want to be regarded as cost centers, providing back-office support. They want to be on the front lines of business, building new and exciting applications that generate new value for their employers. Charting the future of real estate, a sector with myriad impacts on society, can be a meaningful avenue to fulfill this goal.

Integrating data professionals into an organization can be another challenge. To be effective, developers and investors might consider using agile models, originally created for software development. These cross-functional teams can consist of both data and existing investor or development professionals, working collaboratively and reporting to a chief data officer or the chief technology officer. The precise operating model will vary for each organization, but it is critical that stakeholders align on it early and establish the required governance to track impact and generate new ideas.

Many real estate firms have long made decisions based on a combination of intuition and traditional, retrospective data. Today, a host of new variables make it possible to paint more vivid pictures of a location's future risks and opportunities, and with unprecedented granularity. While the technology is still relatively nascent, its predictive power is too great to ignore.

These solutions will evolve at a rapid clip—progress in artificial intelligence is frequently exponential rather than linear—and companies must consider them as realistic supplements to their current underwriting, portfolio review, and research processes. If companies fail to act now, they run the risk of adapting too late. ■

Gabriel Morgan Asaftei is a senior expert in McKinsey's New York office, where **Sudeep Doshi** is an associate partner and Aditya Sanghvi is a partner; **John Means** is a partner in the Washington, DC, office.

This article was originally published in *Urban Land* for the Urban Land Institute.

Designed by Global Editorial Services.
Copyright © 2018 McKinsey & Company.
All rights reserved.